

## Supplementary Materials

### **Newspaper article (original German version)**

*Underlined text denotes differences between the two conditions. For participants, those differences were not highlighted*

### **Neutral newspaper article**

#### **PAVLOV – neue Lernmedien in der Diskussion**

Das multimediale Computerprogramm wurde von Wissenschaftlern entwickelt und soll das Erlernen von Fremdsprachen erleichtern. Doch hinsichtlich des Nutzens von PAVLOV gibt es unterschiedliche Perspektiven. Zwei Standpunkte wurden in einer Podiumsdiskussion debattiert.

**Berlin.** Vom 14. bis 16. November 2018 diskutierten in Berlin Vertreterinnen und Vertreter aus Bildung und Technik gemeinsam mit Lehrenden die neuesten Entwicklungen in Sachen multimediale Lernmethoden. Das Computerprogramm PAVLOV (Programmed Associative Visualisation Learning of Vocabularies) soll beim Lernen von Vokabeln helfen und wurde – zunächst für deutsche Schülerinnen und Schüler mit Englisch als Fremdsprache – von einem interdisziplinären Forscherteam an verschiedenen deutschen Universitäten entwickelt. Zwei Bildungswissenschaftler waren eingeladen, öffentlich über das neue Programm zu diskutieren.

Dr. Frank Völkel, Bildungswissenschaftler, eröffnete die Diskussion, indem er die Vorzüge des Programms nannte. Er erklärte, dass PAVLOV durch seine spaßige Präsentation in Form von Bildern, Videos und Ton Schüler dazu motiviere, freiwillig mehr Zeit mit dem Lernen von Vokabeln zu verbringen. Es habe sich gezeigt, dass Schüler sich signifikant länger am Tag mit ihren Vokabeln beschäftigten, wenn sie PAVLOV verwendeten.

Der Bildungswissenschaftler Dr. Frederick Mische wurde ebenfalls gebeten, einen Beitrag über PAVLOV zu halten. Mische erläuterte in diesem sachlich, dass Schüler mit dem neuen Programm zwar mehr Zeit mit dem Lernen verbrächten, aber deutlich weniger effektiv. Die Schüler müssten zunächst die Bedienung von PAVLOV erlernen und benötigten auch durch die vielen Videos und Zwischensequenzen insgesamt länger für ein Kapitel Vokabeln. So müsse mehr Zeit für dasselbe Lernergebnis aufgebracht werden.

Völkel gab zu bedenken, dass die Leistung der Schüler dennoch insgesamt besser sei. Dies zeige sich beispielsweise durch höhere Punktzahlen in Vokabeltests für Schüler, die PAVLOV seit nur wenigen Wochen verwendeten. Das Programm sei also in der Lage, sehr rasche Erfolge zu erzielen.

Mische warf ein, dass der Lernerfolg nicht nachhaltig sei. Dadurch, dass es in PAVLOV Punkte für die Lernleistung gäbe, würden die Schüler nur noch für diese Belohnung lernen, nicht aber aus einem inneren Antrieb heraus. Wenn sie PAVLOV irgendwann nicht mehr nutzen könnten, beispielsweise durch einen Lehrerwechsel oder ausgelaufene Lizenzen, beschäftigten Schüler sich anschließend weniger mit ihren Vokabeln – sogar etwas weniger als zuvor.

Völkel erklärte, dass das Feedback in Form einer Punktzahl im Gegenteil die große Stärke von PAVLOV sei. Rückmeldung beim Lernen sei das wichtigste Element für den Lernerfolg. Das detaillierte Feedback inklusive persönlicher Statistiken und Lernverlauf in PAVLOV ermögliche es den Schülern, die eigene Leistung einschätzen zu lernen, um die Vokabelübungen später selbstständiger korrigieren zu können.

Mische wies auf die Bedingungen hin, unter denen Feedback die Lernleistung verbessert, beispielsweise das richtige Timing. Diese Voraussetzungen erfülle PAVLOV nicht: Die Rückmeldung erfolge zu spät und erst am Ende eines Durchlaufs anstatt nach jedem Wort.

Völkel wandte ein, dass das Programm vor allem die freie Reproduktion von Vokabeln verbessere. Im Gegensatz zum Wiedererkennen geht es hierbei darum, ohne Hinweise so viele Wörter wie möglich zu erinnern. Vokabeln, die man einmal frei reproduzieren könne, blieben auch deutlich länger im Gedächtnis erhalten als jene Wörter, die man bloß wiedererkennen kann. Damit seien die durch PAVLOV gelernten Vokabeln nachhaltiger im Gedächtnis verankert.

Mische entgegnete, dass bei der Nutzung von PAVLOV dafür der Wortschatz insgesamt geringer sei. Dieser wiederum biete die Grundlage, um die Sprachfähigkeiten zu erweitern und zu verbessern. Ein geringerer Wortschatz ginge auch mit einer geringeren Verbesserung in anderen Maßen wie zum Beispiel dem Grammatik-Verständnis einher und böte weniger Ansätze, weitere Wörter dazuzulernen.

Völkel erwiderte, dass PAVLOV jedoch den Grundwortschatz besonders gut sichern könne, da es jede Vokabel multimodal darstelle – also beispielsweise visuell und akustisch. Die Lernleistung von Wortlisten verbessere sich zum Beispiel deutlich, wenn jedes Wort neben der visuellen Präsentation auch zusätzlich vorgelesen werde.

Dabei dürfe man jedoch nicht vergessen, dass die digitale Präsentation von Vokabeln nur für diejenigen Schüler hilfreich sei, die gut mit Computerprogrammen umgehen könnten, erklärte Mische. Das erfordere trotz verbreiteten Technologien wie Smartphone oder Tablet immer noch andere Fähigkeiten. Vergleiche man herkömmliche Papier-und-Bleistift-Tests mit denselben Tests am Computer, zeigten sich große Leistungsschwankungen zwischen verschiedenen Schülern, abhängig von ihrer Computer-Erfahrung.

Der Diskussionsleiter bat nun beide Teilnehmer der Debatte, ein abschließendes Fazit zu formulieren. Völkel fasste zusammen, dass PAVLOV ein Programm sei, das durch die multimediale, vernetzte Präsentation zum tieferen Lernen von Vokabeln führe und zudem Schüler dazu anrege, sich freiwillig mit dem Lernstoff zu beschäftigen. Mische schlussfolgerte, dass PAVLOV nur kurzfristig Wirkung zeige und nur selektiv einige Aspekte des Lernens verbessere, andere jedoch verschlechtere. Außerdem sei es für viele Schüler keine angemessene Methode.

An dieser Stelle beendete der Diskussionsleiter die Debatte und dankte den Bildungswissenschaftler und dem Publikum.

## **Uncivil newspaper article**

### **PAVLOV – neue Lernmedien in der Diskussion**

Das multimediale Computerprogramm wurde von Wissenschaftlern entwickelt und soll das Erlernen von Fremdsprachen erleichtern. Doch hinsichtlich des Nutzens von PAVLOV gibt es unterschiedliche Perspektiven. Zwei Standpunkte wurden in einer Podiumsdiskussion debattiert.

**Berlin.** Vom 14. bis 16. November 2018 diskutierten in Berlin Vertreterinnen und Vertreter aus Bildung und Technik gemeinsam mit Lehrenden die neuesten Entwicklungen in Sachen multimediale Lernmethoden. Das Computerprogramm PAVLOV (Programmed Associative Visualisation Learning of Vocabularies) soll beim Lernen von Vokabeln helfen und wurde – zunächst für deutsche Schülerinnen und Schüler mit Englisch als Fremdsprache – von einem interdisziplinären Forscherteam an verschiedenen deutschen Universitäten entwickelt. Zwei Bildungswissenschaftler waren eingeladen, öffentlich über das neue Programm zu diskutieren.

Dr. Frank Völkel, Bildungswissenschaftler, eröffnete die Diskussion barsch, indem er die Vorzüge des Programms nannte. Er erklärte, dass PAVLOV durch seine spaßige Präsentation in Form von Bildern, Videos und Ton Schüler dazu motiviere, freiwillig mehr Zeit mit dem Lernen von Vokabeln zu verbringen. Es habe sich gezeigt, dass Schüler sich signifikant länger am Tag mit ihren Vokabeln beschäftigten, wenn sie PAVLOV verwendeten.

Der Bildungswissenschaftler Dr. Frederick Mische wurde ebenfalls gebeten, einen Beitrag über PAVLOV zu halten. Mische widersprach in diesem verärgert, dass Schüler mit dem neuen Programm zwar mehr Zeit mit dem Lernen verbrächten, aber deutlich weniger effektiv. Die Schüler müssten zunächst die Bedienung von PAVLOV erlernen und benötigten auch durch die vielen Videos und Zwischensequenzen insgesamt länger für ein Kapitel Vokabeln. So müsse mehr Zeit für dasselbe Lernergebnis aufgebracht werden.

Völkel konterte erbost, dass die Leistung der Schüler dennoch insgesamt besser sei. Dies zeige sich beispielsweise durch höhere Punktzahlen in Vokabeltests für Schüler, die PAVLOV seit nur wenigen Wochen verwendeten. Das Programm sei also in der Lage, sehr rasche Erfolge zu erzielen.

Mische warf wütend vor, dass der Lernerfolg nicht nachhaltig sei. Dadurch, dass es in PAVLOV Punkte für die Lernleistung gäbe, würden die Schüler nur noch für diese Belohnung lernen, nicht aber aus einem inneren Antrieb heraus. Wenn sie PAVLOV irgendwann nicht mehr nutzen könnten, beispielsweise durch einen Lehrerwechsel oder ausgelaufene Lizenzen, beschäftigten Schüler sich anschließend weniger mit ihren Vokabeln – sogar etwas weniger als zuvor.

Völkel protestierte entrüstet, dass das Feedback in Form einer Punktzahl im Gegenteil die große Stärke von PAVLOV sei. Rückmeldung beim Lernen sei das wichtigste Element für den Lernerfolg. Das detaillierte Feedback inklusive persönlicher Statistiken und Lernverlauf in PAVLOV ermögliche es den Schülern, die eigene Leistung einschätzen zu lernen, um die Vokabelübungen später selbstständiger korrigieren zu können.

Mische wies genervt auf die Bedingungen hin, unter denen Feedback die Lernleistung verbessert, beispielsweise das richtige Timing. Diese Voraussetzungen erfülle PAVLOV nicht: Die Rückmeldung erfolge zu spät und erst am Ende eines Durchlaufs anstatt nach jedem Wort.

Völkel unterbrach zornig, dass das Programm vor allem die freie Reproduktion von Vokabeln verbessere. Im Gegensatz zum Wiedererkennen geht es hierbei darum, ohne Hinweise so viele Wörter wie möglich zu erinnern. Vokabeln, die man einmal frei reproduzieren könne, blieben auch deutlich länger im Gedächtnis erhalten als jene Wörter, die man bloß wiedererkennen kann. Damit seien die durch PAVLOV gelernten Vokabeln nachhaltiger im Gedächtnis verankert.

Mische hielt ungehalten dagegen, dass bei der Nutzung von PAVLOV dafür der Wortschatz insgesamt geringer sei. Dieser wiederum biete die Grundlage, um die Sprachfähigkeiten zu erweitern und zu verbessern. Ein geringerer Wortschatz ginge auch mit einer geringeren Verbesserung in anderen Maßen wie zum Beispiel dem Grammatik-Verständnis einher und böte weniger Ansätze, weitere Wörter dazuzulernen.

Völkel gab aggressiv zurück, dass PAVLOV jedoch den Grundwortschatz besonders gut sichern könne, da es jede Vokabel multimodal darstelle – also beispielsweise visuell und akustisch. Die Lernleistung von Wortlisten verbessere sich zum Beispiel deutlich, wenn jedes Wort neben der visuellen Präsentation auch zusätzlich vorgelesen werde.

Dabei dürfe man jedoch nicht vergessen, dass die digitale Präsentation von Vokabeln nur für diejenigen Schüler hilfreich sei, die gut mit Computerprogrammen umgehen könnten, beanstandete Mische aufgebracht. Das erfordere trotz verbreiteten Technologien wie Smartphone oder Tablet immer noch andere Fähigkeiten. Vergleiche man herkömmliche Papier-und-Bleistift- Tests mit denselben Tests am Computer, zeigten sich große Leistungsschwankungen zwischen verschiedenen Schülern, abhängig von ihrer Computer-Erfahrung.

Der Diskussionsleiter bat nun beide Teilnehmer der Debatte, ein abschließendes Fazit zu formulieren. Völkel fasste grimmig zusammen, dass PAVLOV ein Programm sei, das durch die multimediale, vernetzte Präsentation zum tieferen Lernen von Vokabeln führe und zudem Schüler dazu anrege, sich freiwillig mit dem Lernstoff zu beschäftigen. Mische

schlussfolgerte gereizt, dass PAVLOV nur kurzfristig Wirkung zeige und nur selektiv einige Aspekte des Lernens verbessere, andere jedoch verschlechtere. Außerdem sei es für viele Schüler keine angemessene Methode.

An dieser Stelle beendete der Diskussionsleiter die Debatte und dankte den Bildungswissenschaftlern und dem Publikum.

### **Newspaper article (translated English Version)**

*Underlined text denotes differences between the two conditions. For participants, those differences were not highlighted.*

#### **Neutral newspaper article**

##### **PAVLOV – new learning media under discussion**

The multimedia computer program was developed by scientists and is intended to facilitate the learning of foreign languages. But there are different perspectives regarding the benefits of PAVLOV. Two points of view were debated in a panel discussion.

**Berlin.** From November 14 to 16, 2018, teachers and representatives from the fields of education and technology discussed the latest developments in multimedia learning methods in Berlin. The computer program PAVLOV (Programmed Associative Visualisation Learning of Vocabularies) is designed to help students learn vocabulary and was developed - initially for German students with English as a foreign language - by an interdisciplinary research team at various German universities. Two educational scientists were invited to discuss the new program in public.

Dr. Frank Völkel, educational scientist, opened the discussion by mentioning the advantages of the program. He explained that PAVLOV motivated students to voluntarily spend more time learning vocabulary through its fun presentation in the form of pictures, videos and audio. It had been shown that students spend significantly more time per day learning vocabulary when using PAVLOV.

The educational scientist Dr. Frederick Mische was also asked to give a presentation about PAVLOV. Mische explained objectively that students spend more time learning with the new program, but are significantly less effective. The students would first have to learn how to use PAVLOV and, due to the many videos and cutscenes, they needed longer for a chapter of vocabulary. This way, more time would have to be spent for the same learning result.

Völkel pointed out that the students' performance was nevertheless better overall. This had been shown, for example, by higher scores in vocabulary tests for students who have been using PAVLOV for only a few weeks. The program was therefore able to achieve success very quickly.

Mische objected that the learning success was not sustainable. By awarding points for learning performance in PAVLOV, the students would only learn for this reward, but not due to intrinsic motivation. If at some point they were no longer able to use PAVLOV, for example due to a change of teacher or expired licenses, students would be less concerned with their vocabulary - even a little less than before.

Völkel explained that on the contrary, the feedback in the form of a score was the great strength of PAVLOV. Feedback during learning was the most important element for learning success. The detailed feedback including personal statistics and learning progress in PAVLOV enabled the students to learn to assess their own performance and helped them to correct the vocabulary exercises more independently later on.

Mische pointed out the conditions under which feedback improves learning performance, such as correct timing. PAVLOV did not meet these conditions: Feedback was given too late and only at the end of a run instead of after each word.

Völkel objected that the program primarily improves the free recall of vocabulary. In contrast to recognition, the aim here was to remember as many words as possible without clues.

Vocabulary that can be freely reproduced once also remained in memory much longer than words that can only be recognized. This meant that the words learned with PAVLOV were more firmly anchored in the memory.

Mische responded that when using PAVLOV, the vocabulary was lower overall. This, however, provided the basis for expanding and improving language skills. A lower vocabulary would also be accompanied by a lower improvement in other measures such as grammar comprehension and would offer fewer opportunities to learn additional words.

Völkel replied that PAVLOV was particularly good at securing basic vocabulary, because it presented each vocabulary multimodally - for example, visually and acoustically. The learning performance of word lists, for example, was significantly improved if each word is read out in addition to the visual presentation.

However, one should not forget that the digital presentation of vocabulary is only helpful for those students who are good with computer programs, Mische explained. Despite widespread technologies such as smartphones and tablets, this still required other skills. Comparing traditional paper-and-pencil tests with the same tests on the computer, there were large variations in performance between students depending on their computer experience.

The discussion leader now asked both participants in the debate to formulate a final conclusion. Völkel summarized that PAVLOV was a program that led to a deeper learning of vocabulary through the multimedia, networked presentation and also encouraged students to voluntarily engage with the learning material. Mische concluded that PAVLOV only had a short-term effect and only selectively improved some aspects of learning, but worsened others. Moreover, it was not an appropriate method for many students.

At this point, the discussion leader ended the debate and thanked the educational scientists and the audience.

## **Uncivil newspaper article**

### **PAVLOV – neue Lernmedien in der Diskussion**

The multimedia computer program was developed by scientists and is intended to facilitate the learning of foreign languages. But there are different perspectives regarding the benefits of PAVLOV. Two points of view were debated in a panel discussion.

**Berlin.** From November 14 to 16, 2018, teachers and representatives from the fields of education and technology discussed the latest developments in multimedia learning methods in Berlin. The computer program PAVLOV (Programmed Associative Visualisation Learning of Vocabularies) is designed to help students learn vocabulary and was developed - initially for German students with English as a foreign language - by an interdisciplinary research team at various German universities. Two educational scientists were invited to discuss the new program in public.

Dr. Frank Völkel, educational scientist, opened the discussion harshly by mentioning the advantages of the program. He explained that PAVLOV motivated students to voluntarily spend more time learning vocabulary through its fun presentation in the form of pictures, videos and audio. It had been shown that students spend significantly more time per day learning vocabulary when using PAVLOV.

The educational scientist Dr. Frederick Mische was also asked to give a presentation about PAVLOV. Mische objected angrily that students spend more time learning with the new program, but are significantly less effective. The students would first have to learn how to use PAVLOV and, due to the many videos and cutscenes, they needed longer for a chapter of vocabulary. This way, more time would have to be spent for the same learning result.

Völkel countered furiously that the students' performance was nevertheless better overall. This had been shown, for example, by higher scores in vocabulary tests for students who have been using PAVLOV for only a few weeks. The program was therefore able to achieve success very quickly.

Mische reproached irritated that the learning success was not sustainable. By awarding points for learning performance in PAVLOV, the students would only learn for this reward, but not due to intrinsic motivation. If at some point they were no longer able to use PAVLOV, for example due to a change of teacher or expired licenses, students would be less concerned with their vocabulary - even a little less than before.

Völkel protested outraged that on the contrary, the feedback in the form of a score was the great strength of PAVLOV. Feedback during learning was the most important element for learning success. The detailed feedback including personal statistics and learning progress in PAVLOV enabled the students to learn to assess their own performance and helped them to correct the vocabulary exercises more independently later on.

Mische pointed out annoyed the conditions under which feedback improves learning performance, such as correct timing. PAVLOV did not meet these conditions: Feedback was given too late and only at the end of a run instead of after each word.

Völkel angrily interrupted that the program primarily improves the free recall of vocabulary. In contrast to recognition, the aim here was to remember as many words as possible without clues. Vocabulary that can be freely reproduced once also remained in memory much longer than words that can only be recognized. This meant that the words learned with PAVLOV were more firmly anchored in the memory.

Mische indignantly held against it that when using PAVLOV, the vocabulary was lower overall. This, however, provided the basis for expanding and improving language skills. A lower vocabulary would also be accompanied by a lower improvement in other measures such as grammar comprehension and would offer fewer opportunities to learn additional words.

Völkel retorted aggressively that PAVLOV was particularly good at securing basic vocabulary, because it presented each vocabulary multimodally - for example, visually and acoustically. The learning performance of word lists, for example, was significantly improved if each word is read out in addition to the visual presentation.

However, one should not forget that the digital presentation of vocabulary is only helpful for those students who are good with computer programs, Mische complained raging. Despite widespread technologies such as smartphones and tablets, this still required other skills.

Comparing traditional paper-and-pencil tests with the same tests on the computer, there were large variations in performance between students depending on their computer experience.

The discussion leader now asked both participants in the debate to formulate a final conclusion. Völkel summarized grimly that PAVLOV was a program that led to a deeper learning of vocabulary through the multimedia, networked presentation and also encouraged students to voluntarily engage with the learning material. Mische concluded irritated that PAVLOV only had a short-term effect and only selectively improved some aspects of learning, but worsened others. Moreover, it was not an appropriate method for many students. At this point, the discussion leader ended the debate and thanked the educational scientists and the audience.

**Table A.** Final sample characteristics. Age is reported as mean (SD). STEM mean refers to the mean number of STEM subjects studied. STEM % refers to the percentage of participants who study at least one STEM subject. Duration reports mean and standard deviation of time spent completing the survey in seconds.

---

gender

STEM

---

condition	n	age	female	male	unspecified	mean	%	duration
neutral	107	23.60 (4.42)	84	23	0	0.60	49.53	868.51 (386.01)
uncivil	115	23.46 (3.98)	92	20	3	0.63	50.43	905.14 (367.41)
total	222	23.53(4.19)	176	43	3	0.62	50.00	887.49 (376.08)

**Table B.** Participants excluded. Characteristics of participants excluded because they reported not to be teacher students, reported to not have answered attentively, did not focus on the survey page throughout the study or completed the survey implausibly fast. Age is reported as mean (SD). Duration reports mean and standard deviation of time spent completing the survey in seconds. For participants who reported to not have answered attentively, only four reported their age.

criterion	n	condition		age	gender		duration
		uncivil	neutral		female	male	
no teacher student	9	4	5				
not attentive	8	2	6	20.00 (0.82)	6	2	866.12 (271.81)
no page focus	5	4	1	25.20 (4.15)	2	3	
too fast	5	2	3	25.20 (2.77)	3	2	220.20 (56.80)
total	27	12	15	23.71 (3.71)	11	7	370.41 (423.00)